CENTENNIAL PAPER: Proteomics in animal science

J. D. Lippolis and T. A. Reinhardt

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://jas.fass.org/cgi/content/full/86/9/2430

American Society of Animal Science

www.asas.org

CENTENNIAL PAPER: Proteomics in animal science

J. D. Lippolis[1] and T. A. Reinhardt

Periparturient Diseases of Cattle Research Unit, USDA-ARS, National Animal Disease Center, Ames, IA 50010

ABSTRACT: Proteomics holds significant promise as a method for advancing animal science research. The use of this technology in animal science is still in its infancy. The ability of proteomics to simultaneously identify and quantify potentially thousands of proteins is unparalleled. In this review, we will discuss basic principles of doing a proteomic experiment. In addition, challenges and limitations of proteomics will be considered, stressing those that are unique to animal sciences. The current proteomic research in animal sciences will be discussed, and the potential uses for this technology will be highlighted.

INTRODUCTION

It seems part of our nature as a species to manipulate our surroundings. This includes the animals that we use for food and companionship. Although genetic manipulation in food animals is controversial, it has been going on for centuries as we have bred animals for certain characteristics. Today, we still manipulate animal characteristics by breeding for qualities of interest. Detailed records of animals used for breeding are kept and their progeny observed to find those animals with the best production results. Thus, we have leaner animals for consumption or dairy animals that produce 100 pounds of milk a day. With a breeding approach to animal husbandry we can predict which animal will yield offspring with a desired trait. Breeding for a trait can be done with little or no knowledge of which genes are responsible for the desired trait. The genetic knowledge becomes important when linkage to a disease is found in a specific breeder. For example, bovine leukocyte adhesion deficiency (**BLAD**) is an autosomal recessive genetic condition that is lethal. It was shown that BLAD was due to a single nucleotide mutation in the gene CD18. Using PCR, researchers were able to determine which sires had the recessive gene, enabling their removal from the breeding population (Ryncarz et al., 1995). Thus, knowledge of the gene and its sequence was critical to the identification and resolution of this disease. With the advent of the genomic age we are moving away from breeding for a physical trait and moving toward breeding for a particular gene or set of genes. In fact, genetically modified cows were recently generated that produced an antibacterial protein in their milk resulting in animals less susceptible to *Staphylococcus aureus* mastitis (Wall et al., 2005). The genome projects of the various domestic animals will increase our ability to associate desired traits with the necessary genes. However, understanding the genes is only the beginning of the biological story.

Genes are transcribed into mRNA, mRNA is translated into proteins, and proteins form and build each cell. Therefore, the gene sequence, transcription, translation, and posttranslation are all points in a gene product's expression that are controlled. Genome projects have given us the gene sequence information, techniques such as real-time PCR and microarray are giving us gene expression data, and proteomics is giving us information regarding protein expression and posttranslational modifications. Together these techniques have brought animal research to a molecular level.

In addition to mutations in genes (e.g., CD18, see above), the expression of genes can be very important in disease manifestation. There are 2 ways to determine gene expression: first, quantitation of mRNA, and second, detection or relative quantitation of protein. Both measurement of mRNA and protein can be accomplished using techniques that measure single genes (e.g., real-time PCR for mRNA and Western blot for protein) or genome-wide technologies that measure thousands of genes (e.g., microarray for mRNA and proteomics for proteins). Microarray analysis is used to determine mRNA expression for as many as 10,000 to 40,000 different genes (Stoughton, 2005; Ness, 2007). Microarray analysis has the advantage of being able to

measure gene expression changes in a large percentage of the genes of a cell in one experiment. In contrast to proteomic studies that suffer from dynamic range issues (i.e., high abundance proteins overwhelming detection of low abundance proteins), microarray can detect not only mRNA from high expressed genes but also very low expressed genes. However, numerous studies have shown examples of a lack of correlation between mRNA and protein abundance (Gygi et al., 1999b; Ideker et al., 2001; Griffin et al., 2002). For example, changing the carbon source for yeast resulted in a 500-fold increase in mRNA for a gene involved in sugar metabolism, whereas the corresponding protein only increased 10-fold. In addition, some genes showed no change in mRNA levels but showed significant increases in protein levels (Griffin et al., 2002). These examples of a lack of correlation highlight the importance of linking mRNA expression results with subsequent proteomic studies.

Proteomics is the large-scale study of protein expression, protein-protein interactions, or posttranslational modifications (for more specific reviews, see Ong and Mann, 2005; Cravatt et al., 2007; Gingras et al., 2007; Witze et al., 2007). Unlike other methodologies that analyze a few proteins at a time, proteomics can analyze thousands of proteins in a single experiment. This ability to analyze thousands of proteins gives the field of proteomics a unique capability to demonstrate how cells can dynamically respond to changes in their environment. Therefore, a goal of proteomics is to identify new and potentially unexpected changes in protein expression, interaction, or modification as a result of an experimental treatment. Generation of large proteomic data sets is expected to demonstrate the interdependence of cellular processes important for normal cell growth or a cell's response to abnormal or disease conditions. In essence, a proteomic approach enables an investigator to step back and, without prejudice, view the whole picture of cellular functions instead of one particular action of one protein. This type of research enables the discovery of unexpected connections between cellular processes as a precursor to new hypotheses.

In this review we will discuss some of the current proteomic technologies and challenges of these technologies. In addition, we will discuss how the completion of the Human Genome Project has helped propel proteomics research and how genome projects in various agricultural species will aid proteomic studies in animal science. Finally, we will discuss how these technologies are being applied to solve important scientific questions in agricultural animals and animal products.

PROTEOMICS

The field of proteomics exists because of the advances in the last decade in the fields of mass spectrometry, genomics, and bioinformatics. Both sensitive instrumentation and a complete protein database are essential to most proteomic work. Advancements in mass spectrometer design are predicted to increase the number of proteins identified, with the goal of an accurate and complete list of proteins that constitute a proteome. Continued refinement and annotation of genomic databases will enable identification and functional assignment of mass spectrometry data.

Examples of Experimental Approach

Proteomic experiments can be directed toward detection of certain known proteins of interest, or an indirect or shotgun approach can be taken. In the shotgun approach large complex proteomes are analyzed for changes in protein presence, expression, or modification. Generation of a large data set of protein expression can be an important prelude to hypothesis-driven research and can yield unexpected results. The strength of this technology is the sensitivity of the mass spectrometer and the number of proteins that this type of experiment can identify. Whereas procedures such as Western blot detect one protein per experiment, a single proteomic experiment can identify greater than a thousand proteins. In addition, techniques such as Western blot require reagents such as a specific antibody for detection of a protein; shotgun proteomics requires no a priori knowledge of a protein's existence in a sample.

Technical advances have improved the sensitivity and accuracy of mass spectrometers necessary for proteomic work. Despite this high sensitivity, 2 factors complicate protein identification: first, the number of proteins that constitute a proteome, and second, the expression level range. First, the number of proteins that constitute the human proteome is estimated to be greater than 30,000 proteins, not counting alternative splice variants and posttranslational modifications (Cho, 2007). Second, the range of protein expression complicates detection of low abundance proteins in typical biological samples. The expression dynamic range is estimated to be greater than 7 orders of magnitude (Stasyk and Huber, 2004). For example, nearly half of the protein in plasma is albumin, and the top 10 proteins in plasma make up nearly 90% of the total protein (Cho, 2007). These 2 factors have led to the addition of various protein separation methods to proteomic experiments before mass spectrometry. Therefore, various fractionation schemes of the proteome into less complex mixtures are necessary for a more complete identification of proteins. Fractionation can be achieved by subcellular fractionation, enrichment strategies, chromatography, or gel electrophoresis (Stasyk and Huber, 2004). These fractionation strategies can be used individually or in combination to improve detection of small abundance proteins.

Figure 1 shows an outline of a potential proteomic experiment with multiple dimensions of sample fractionations. After sample isolation, one of the first potential

fractionation methods is subcellular fractionation of the cell (Warnock et al., 2004). Use of this method has yielded information such as the proteomic composition of the nuclear membrane. Interestingly, protein composition of the nuclear membrane is unique in different tissues or different points in development (Schirmer and Gerace, 2005). For example, isolation of the milk fat globule membrane (**MFGM**) from bovine milk is thought to represent the apical membrane of secretory epithelial cells. Proteomic analysis of the MFGM could enhance the understanding of the molecular mechanisms of lactation (Reinhardt and Lippolis, 2006). In addition, the proteomic analysis provided evidence that these cells express toll-like receptors, which may indicate their participation in an immune response to a mammary gland infection (Reinhardt and Lippolis, 2006). Isolation of specific cellular organelles not only serves as a first step fractionation to reduce the complexity of the proteome, but could also lead to discoveries regarding the nature and functions of proteins resident to the organelle.

There are several fractionation schemes that can separate proteins or peptides by their various physical attributes subsequent to or in place of subcellular fractionation (Figure 1). These schemes may include enrichment of proteins using molecules with specific binding affinities to isolate away a group of proteins, gel electrophoresis, or various chromatographies.

Antibodies and metal ions are 2 common means to enrich for a specific type of protein. Antibodies specific for a class of molecules called major histocompatibility complex (**MHC**) have been used to isolate these molecules away from other cellular proteins. The MHC molecules bind to peptide fragments of proteins normally from the cell and importantly from pathogens, and then present the antigen on the cell surface for detection by the immune system. The peptides are separated from the proteins using size exclusion filtration, and sequenced using mass spectrometry (Hunt et al., 1992; Lippolis et al., 2002). For example, this type of proteomic approach has yielded information regarding the nature of the autoantigens that may be involved in the autoimmune disease type 1 diabetes (Nepom et al., 2001). Both antibodies and metal ions (immobilized metal ion affinity chromatography) have been used to enrich for phosphorylated proteins (Ptacek and Snyder, 2006). It is estimated that 30% of cellular proteins are phosphorylated, and phosphorylation often acts as an on/off switch for the protein's function.

There are several types of gel electrophoresis procedures that are common to proteomic experiments. The most common type of gel electrophoresis associated with proteomic experiment is 2-dimensional SDS-PAGE (**2D-PAGE**). This type of gel electrophoresis separates proteins by 2 different physical attributes (Stasyk and Huber, 2004; Elrick et al., 2006). First, proteins are separated by their isoelectric points in an isoelectric focusing (**IEF**) gel. These IEF gels are made with broad or narrow pH ranges to allow for different levels of resolution. After proteins have migrated in the gel to their isoelectric point, the gel is then placed orthogonally to a typical reducing SDS-PAGE gel, and proteins are run into that gel and separated on the basis of molecular weight (Figure 1, see example of 2D-SDS-PAGE gel). The results are hundreds of protein spots on a gel where left-right separation is by IEF and up-down separation is by SDS-PAGE. Each protein spot can then be cut out and the proteins analyzed by mass spectrometry. Similar to 2D-PAGE is blue native gel electrophoresis (**BN-PAGE**) that replaces the first dimension IEF with a nondenaturing electrophoresis. The BN-PAGE will allow protein complexes to remain intact during the first dimension and separate them by the molecular weight of the complexes. The first dimension native gel is then placed orthogonally to an SDS-PAGE gel, which then separates the complexes into the constituent proteins by their individual molecular weights (Reisinger and Eichacker, 2006). Thus, protein complexes can be analyzed and constituent protein identified using BN-PAGE technology. One-dimensional SDS-PAGE is also an important tool in protein separation before mass spectrometry. Because of its broad molecular resolving range, one-dimensional-PAGE can provide a convenient and powerful means to separate protein for proteomic studies (Lippolis and Reinhardt, 2005; Reinhardt and Lippolis, 2006).

The most common form of peptide separation in proteomic research is liquid chromatography. Peptides can be separated into fractions based on their hydrophobicity in the case of reverse-phase chromatography (**RP-HPLC**) or based on their charge state in the case of strong cation exchange chromatography. Capillary or nanobore columns (25 to 100 μm) with flow rates less than 200 nL/min and the volatile nature of the solvents used in RP-HPLC allows this type of chromatography to be directly linked to mass spectrometers (**LC-MS**). Therefore, peptide separation and peptide detection can be simultaneous. Other fractionation techniques can be used upstream of LC-MS to provide greater separation. In what is commonly referred to as multidimensional protein identification technology (**MudPIT**) 2 liquid chromatography techniques are used in tandem (Washburn et al., 2001). Linking strong cation exchange chromatography and LC-MS, researchers were able to detect nearly 1,500 proteins from a yeast proteome (Washburn et al., 2001).

The goal of these various protein or peptide separation schemes is to enable detection of all the proteins in a proteome. When chromatography is linked to the mass spectrometer, peptides will elute from the column in finite peaks. A mass spectrometer can sample and sequence only one peptide at a time (approximately 1 peptide every 1.5 s). Many peptides may be missed if the sample is too complex. For example, if a whole proteome (30,000 proteins) is digested into 10 peptides each and was separated evenly over a 1-h chromatography gradient, the average number of peptides eluting off the column would be over 80 peptides per second.
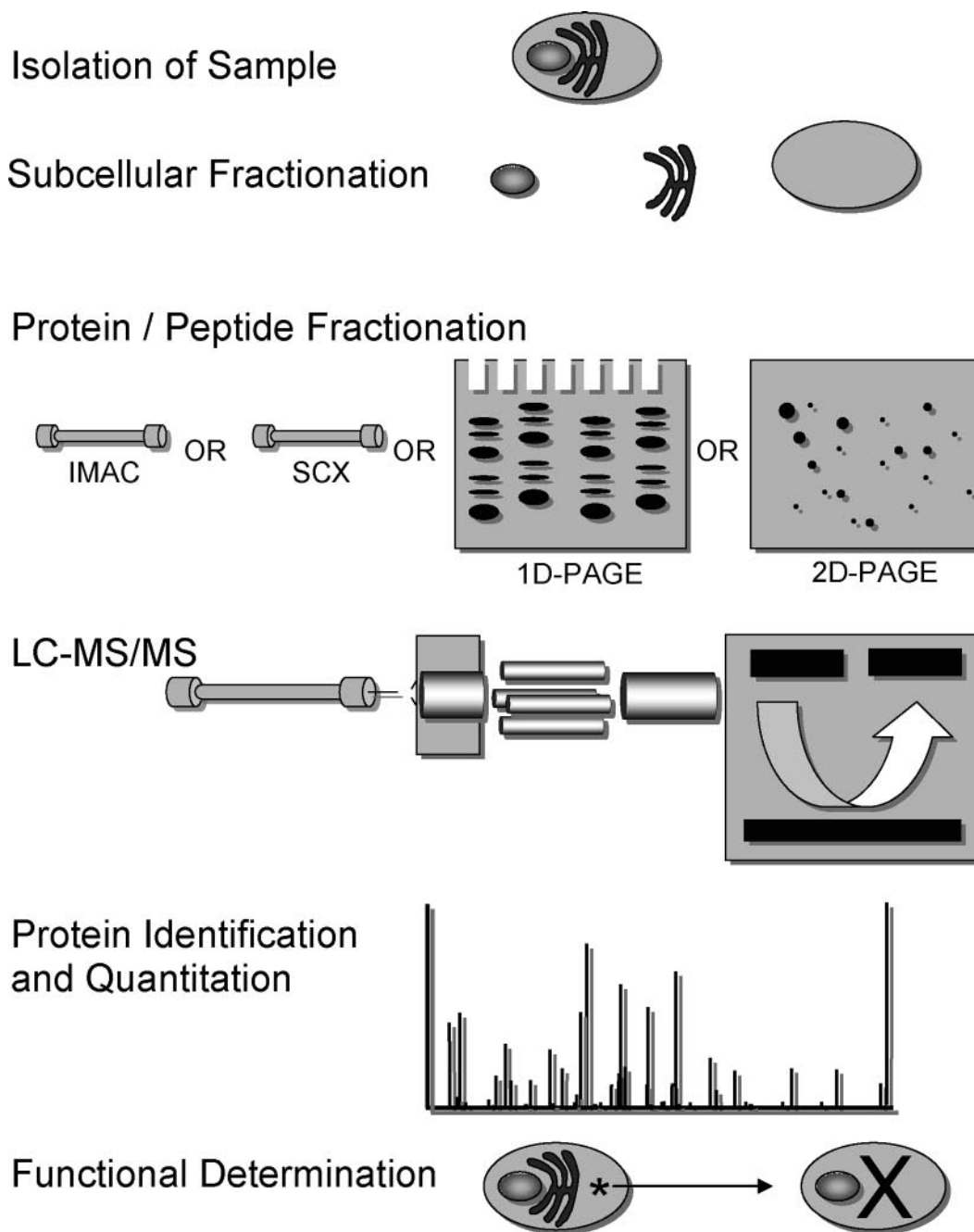
Figure 1. General proteomic experimental design. Multiple separation methodologies can be employed in a proteomic experiment to maximize the number of proteins identified. This figure illustrates some potential separations that may be used singly or in combination. IMAC = immobilized metal ion affinity chromatography. SCX = strong cation exchange chromatography. 1D-PAGE = 1-dimensional PAGE. 2-D-PAGE = 2-dimensional PAGE. LC-MS/MS = liquid chromatography linked to a tandem mass spectrometer.

For this reason, multiple separations are necessary to reduce the number of peptides per second introduced into the mass spectrometer for the analysis of as many peptides as possible.

Mass Spectrometry

Current mass spectrometers can detect and identify peptides in the femtomole (10^{-15}) to attomole (10^{-18}) range (Moyer et al., 2003). There are many types of mass spectrometers that can be used for proteomic studies, and each accomplishes the task of peptide identification in a slightly different way. However, the basic process of identifying a protein using a mass spectrometer is consistent between the various types. After initial protein digestion typically with trypsin, peptides must be ionized to enter the mass spectrometer. Peptides are then detected, isolated, and finally, fragmented and sequenced by the mass spectrometer.
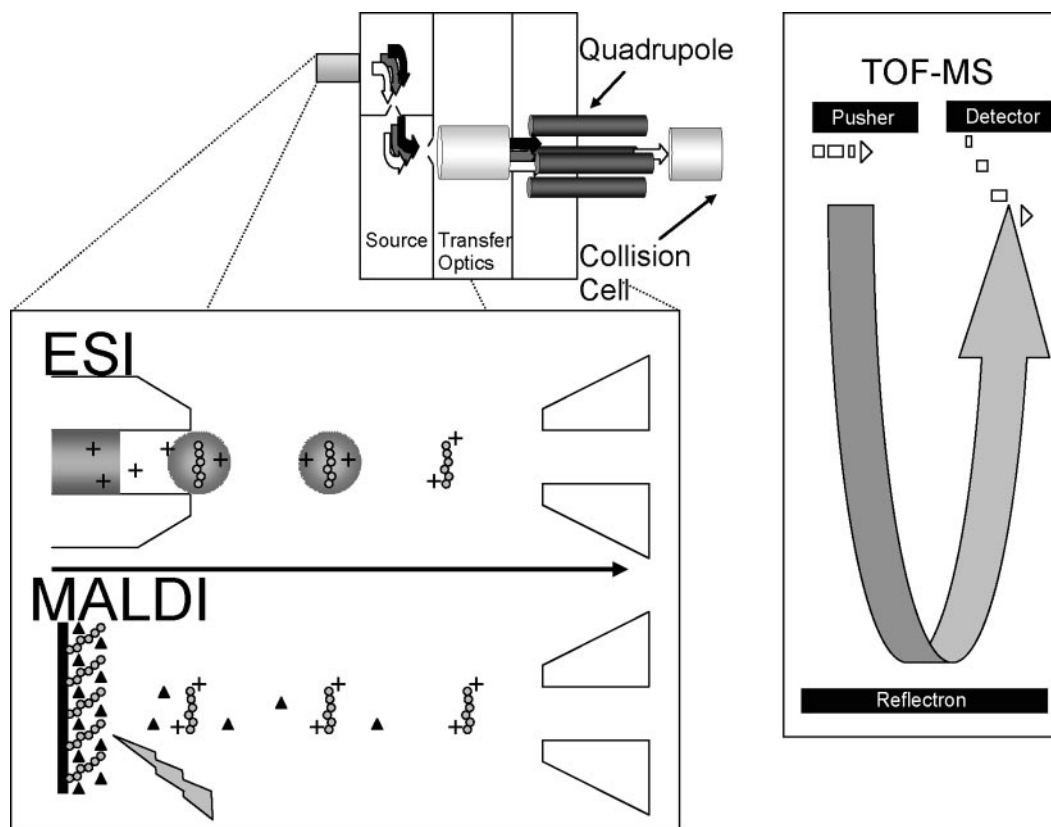
Figure 2. Outline of tandem mass spectrometry. Ionization of peptides is accomplished by one of 2 methods: electrospray ionization (ESI) or matrix-assisted laser desorption ionization (MALDI). Charged gas-phase peptides are generated by ESI or MALDI when the acidic HPLC solution containing peptides evaporates or sublimation of peptides occurs when a peptide-crystal matrix is hit with a laser. Ionized gas-phase peptides are then drawn into the mass spectrometer. The qTOF is a tandem mass spectrometer that has a quadrupole mass analyzer in front of a time of flight (TOF) mass analyzer. Ionized peptides travel in a constant stream through the instrument to the TOF. There, distinct packets of peptides are pushed orthogonally to their original flight path. A mass spectrum (MS) containing all the peptides in that package can be observed. To sequence a peptide, the quadrupole is automatically set to allow only a single peptide to pass. The single peptide is then fragmented in the collision cell before entering the TOF, and the peptides fragments are analyzed.

Ionization of peptides is the first step in mass spectrometry of proteomes. The 2 frequently used ionization methods are electrospray ionization (**ESI**) and matrix-assisted laser desorption ionization (**MALDI**; Yates, 1998). One advantage of ESI is that this method of ionization allows for the direct linkage between liquid chromatography and mass spectrometry because of the volatility of the HPLC solutions. Charged gas-phase peptides are generated by ESI when the acidic HPLC solution containing peptides is sprayed from a tip, and the solution evaporates (Figure 2). The MALDI requires mixing of the peptide with a UV-absorbing molecule and the formation of crystals. When a laser strikes the crystalline structure, the results are the sublimation of the matrix and ionization and release of the associated peptides (Figure 2). The peptides are then analyzed by the mass spectrometer and the peptide mass determined. The peptide's mass is typically expressed as a ratio of mass divided by the charge of the peptide (m/z). Both ESI and MALDI cause peptides to gain protons. The same peptide population may gain

a different number of protons. Therefore, a peptide with a mass of 1,000 Da will be detected on a mass spectrometer with an m/z of 1001 if it gained 1 proton, 501 if it gained 2 protons, and 334.3 if it gained 3 protons during ionization.

There are 2 basic mass spectrometry methods for the identification of proteins. The first called peptide fingerprinting is often associated with 2D-PAGE protein separation scheme. Individual spots from the 2D-PAGE gel are isolated and the proteins digested with typsin. These proteins are typically analyzed with a MALDI-time of flight mass spectrometer. The mass spectrometer will record all the peptide m/z detected in the gel spot. Identification of a protein is based on the measurement of multiple peptides that can come from that protein. For example, after a mass spectrometer has determined the m/z for the peptides from a gel spot, this information will be matched to a protein database. A successful protein match will be based on the number of peptides matched to the protein and the accuracy of the matches. The second method to identify proteins

involves use of tandem mass spectrometers that allow for sequencing of individual peptides.

There are many types of mass analyzers that are appropriate for proteomics. Each has strengths and weaknesses. Many times 2 mass analyzers are used in tandem, allowing sequencing of peptides. For example, the qTOF mass spectrometer has a quadrupole (q) mass analyzer in front of a time of flight (**TOF**) mass analyzer (Figure 2). In this example peptide analysis occurs in 2 steps. The first step to peptide analysis is the detection of the group of peptides being ionized at any one time (Figure 3). After peptides are ionized, a constant stream of peptides travels through the instrument to the TOF. There, distinct packets of peptides are pushed orthogonally to their original flight path. The time that it takes each peptide in that packet to travel to the detector is proportional to its mass divided by its charge (m/z). A mass spectrum (**MS**) containing all the peptides in that package can be observed (Figure 3). The second step in the analysis of the peptide sample then occurs. The operating software, typically based on peptide abundance, selects dominant ions from the MS spectra for peptide sequencing. The function of the quadrupole in a qTOF instrument is to allow only one ion of choice to pass through. The single peptide is fragmented before entering the TOF, and packets of the peptides fragments are analyzed. Interpretation of the resulting spectrum (**MSMS**) allows for the sequencing of the peptide (Figure 3). The different formats of mass spectrometers accomplish the task of analyzing and sequencing peptides in different ways. However, each type of mass spectrometer detects the composition of a group of peptides, isolates a specific peptide, fragments the specific peptide, and finally, detects the peptide fragments.

Quantitation

Mass spectrometry not only offers a means to identify a large number of proteins, but also can be used to determine expression levels of proteins. For example, a quantitation method allowed the comparison of the bovine neutrophil proteome taken from dairy cattle weeks before calving with the neutrophil proteome taken at the time of calving, when neutrophils are known to be functionally suppressed. Investigators showed over 40 proteins were differentially expressed between normal and immunosuppressed neutrophils (Lippolis et al., 2006a). Many of these proteins had not been previously associated with immunosuppression.

There are 3 commonly used techniques to quantify proteins in a large-scale proteomic experiment. Two of the techniques use molecular tags that label peptides through an interaction with a free thiol or amine group, and the other technique employs 2D-PAGE gels to determine changes in protein expression. The first molecular tags called isotope-coded affinity tags (**ICAT**) came as a paired set; one tags with a light tag and the other with a heavy tag. Using ^{13}C atoms, the heavy tag

is 9 Da heavier than the light tag and yet maintained the same chromatography profile. Comparison of the proteome of a cell type grown in 2 conditions is analyzed by labeling one sample with the light tag and labeling the other sample with the heavy tag. The 2 samples are mixed and run on the reverse-phase chromatography at the same time. Mass spectrometry analysis of the abundance of 2 co-eluting peptides whose masses were different by 9 Da would identify a pair of peptides, one from each sample group (Gygi et al., 1999a). A newer abundance tag system called iTRAQ uses a set of 4 tags that have the same molecular weight but differ in their fragmentation points. With this protocol, 4 unique samples can be analyzed simultaneously, each labeled with a unique tag. The samples are mixed after labeling. Labeling has no effect on the chromatography profiles of the peptides. Therefore, the same peptide from the 4 samples will comigrate through chromatography and enter the mass spectrometer at the same time. Because the tags have the same mass, they will be selected for MSMS sequencing by the mass spectrometer as a group. Fragmentation of the peptides in MSMS not only cause the breakup of the peptide in a way that is used for peptide sequencing but also causes the breakup of the tag. Upon fragmentation each iTRAQ tag releases a unique fragment; the abundance of the unique fragment is proportional to the abundance of the peptide from a specific sample (Ross et al., 2004).

Running parallel 2D-PAGE is another common method to determine protein expression changes. This approach requires replicate runs to overcome variations in running the gels to gain confidence. Using various protein dyes, the intensity of the protein spots are compared and using mass spectrometry the identity of proteins determined. A variation to 2D-PAGE has been developed and referred to as difference gel electrophoresis (**DIGE**). In DIGE, protein dyes called CyDye fluors are used to label proteins before 2D-PAGE. These dyes are matched for mass and charge and therefore do not greatly alter protein migration in the 2-dimensional gel. Consequently, if 2 samples are labeled with different CyDyes, the same protein from both samples will comigrate in the 2D-PAGE. Fluorescence characteristics of the 2 dyes would be used to determine the relative abundance of that protein from the 2 samples (Wu et al., 2000).

Comparisons of these expression proteomics methods have shown the strengths and weaknesses of each (Wu et al., 2000). A difference between ICAT and iTRAQ is that ICAT labels peptides through a specific interaction with cysteine residues, whereas iTRAQ labels peptides through an interaction with amine groups. The result is that ICAT will only label a fraction of the peptides in the population, whereas the iTRAQ will label all peptides. The result is that iTRAQ will label more peptides per protein, and thus more expression data points will be collected. However, ICAT labeling occurs before trypsin digestion, whereas iTRAQ labeling occurs after. Greater handling of sample before labeling
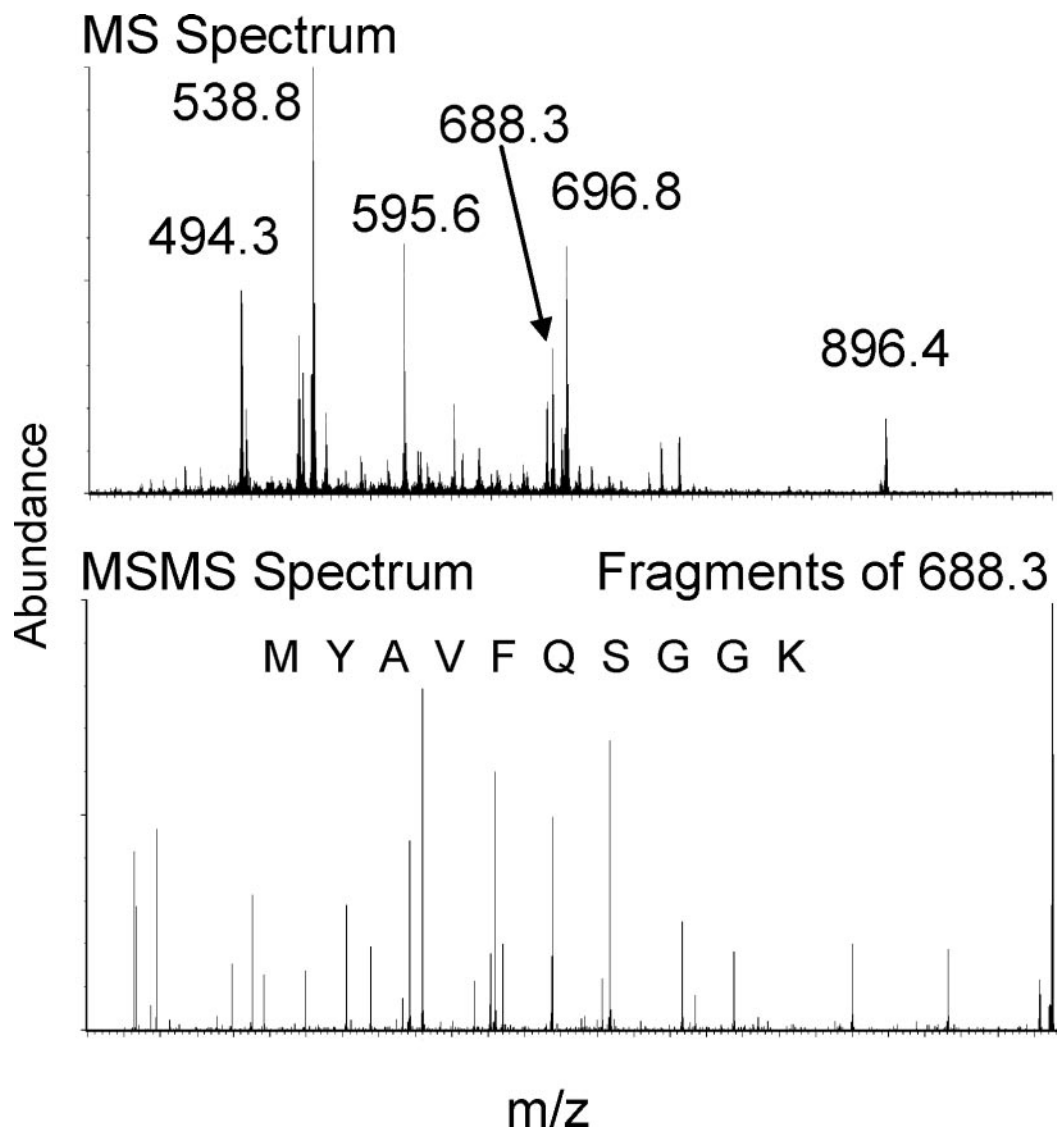
Figure 3. Mass spectra: mass spectrometry (MS) and tandem mass spectrometry (MSMS) data. These data represent MS and MSMS data from a tandem mass spectrometer with a quadrupole mass analyzer in front of a time of flight mass analyzer. The top data panel contains a sample MS spectrum. This represents a single 1.5-s scan of the instrument in a 2 h experimental run. At this time point many peptides can be observed. The instrument computer software will select ions with sufficient abundance for subsequent MSMS analysis for sequencing. The bottom panel contains information that led to the sequencing of the 688.3 peptide seen in the top panel.

could be a source of experimental error and consistent sample handling must be ensured. Disadvantages of the DIGE systems were that in one study more than one-third of the protein spots contained more than one protein, in one case as many as 6 (Wu et al., 2000). Comigration of several proteins will compromise the accuracy of the quantitation using the DIGE methodology. The conclusion of Wu and coworkers was that ICAT, iTRAQ, and DIGE were reasonably good methods for quantitative proteomics but that iTRAQ is the most sensitive (Wu et al., 2000).

Bioinformatics

A proteomic experiment that includes multiple dimensions of fractionation before mass spectrometry can lead to the generation of tens of thousands of peptide MSMS spectra. Each spectrum must be analyzed and a sequence of the peptide determined. Various software has been generated to either de novo sequence or match the spectra to known peptides. These peptides sequences are associated to proteins, and the proteins are identified with variable confidences based on the number and quality of the peptides sequenced. Thus, tens of thousands of peptide spectra can be analyzed and distilled into a report of hundreds of proteins in minutes. The challenge for researchers who use proteomics as a tool is how to interpret the data. When is one confident in the protein identification? How many unique peptides are required to identify a protein? What is the false discovery rate? How are closely related proteins distinguished? These questions are presently a matter

of debate in the proteomics field. However, these are the questions that one must address when writing and reading a proteomics paper.

Most of the protein identification software (e.g., MASCOT, Seaquest) starts with the assumption that the peptide that is being sequenced is in the protein database. If this assumption is not true, these programs will find the best matching peptide in the database. In the case of highly homologous proteins, sufficient identity may exist to correctly sequence a peptide and make a protein match in another closely related species. For example, a proteomic study of an agricultural animal tissue may report protein matches to animals that have extensive protein databases (e.g., human, mouse, or rat). This is likely due to incomplete nature of the current agricultural animal protein database compared with these other animal protein databases. In addition, these protein sequencing software packages allow for several mismatches between the mass spectrometry data and the database; this allows identification of closely related proteins but can affect the confidence scoring of the protein. Thus, the quantity and quality of protein identification is affected by the quality of the protein databases.

Proteomics owes its existence, in part, to the Human Genome Project. Completed in 2003, this 13-year project had as 2 of its goals to identify all of the approximately 20,000 to 25,000 genes in human DNA and determine the sequence of the 3 billion chemical base pairs that make up human DNA. The success of this project has resulted in sequencing projects in other model species (e.g., mouse and rat), companion species (e.g., dog and cat), and agricultural species (e.g., cow, pig, horse, and sheep). However, the genomic databases of these species are not at the same level of completeness. For example, the number of genes currently identified in the NCBI database is 64,000 for mouse, 29,000 for bovine, and 3,000 for pig. As sequencing projects for these various species are completed, the number and confidence of proteomic techniques to identify proteins will be increased. For example, Figure 4 shows the effect on protein identification using the protein database from 2 different years. The same mass spectrometry data set was processed using a database from October 2004 or August 2007. Importantly, in August of 2006 the Bovine Genome Sequencing Center released an updated version of the genomic assembly, adding more genomic information to the bovine protein database. In Figure 4A the data show that 15.6% of the proteins identified using the 2004 database were bovine, compared with 30.6% using the 2007 database. Not only did the number of proteins identified as bovine increase, but the average protein score increased, indicating greater confidence in the identification (Figure 4B). From these data we conclude that a more complete database (the 2007 compared with the 2004) results in a better data set as determined by the number of proteins of the correct species and the higher identification scores. Therefore, as sequencing projects for various agricultural species are completed and annotated, the quality and quantity of protein identifications will increase. This highlights the need for continued effort into complete sequencing of domestic animal genomes.

The presence, absence, or modification of a protein has limited value without knowledge of the function of the protein. The modulation of a protein with a known function can then be associated with a cellular process and the change in a cellular process linked to perturbation of the normal homeostasis at a cellular level. The challenge of proteomics is to sort through a mountain of data and find the information about protein changes that are critical to the host's response to changes in its environment or disease. Efforts to categorize proteins according to their known or predicted functions are currently underway. Consortiums such as The Gene Ontology (http://www.geneontology.org) and the Kyoto Encyclopedia of Gene and Genomes (http://www.genome.jp/kegg) provide means to group proteins by function or into biological processes.

CURRENT STATE OF PROTEOMIC RESEARCH IN ANIMAL SCIENCES

The application of mass spectrometer-based proteomics to agriculturally relevant animal physiology, immunology, reproduction, muscle, and lactational biology is in its infancy and still limited in scope. We will highlight some relevant proteomic work in the animal sciences to date.

Proteomics in animal sciences can be used to survey the proteins expressed in cells, cell compartments, tissues and fluids, which yield data libraries of the abundant proteins (Lippolis and Reinhardt, 2005; Cho et al., 2006; Reinhardt and Lippolis, 2006; Radosevich et al., 2007). Some of these new protein finds should lead to future hypothesis-driven research. Quantitative proteomics can be used to identify a protein and its expression in samples from different experimental treatments; diseased versus healthy animals and disease-causing bacteria at the site of disease versus the same bacteria grown in the laboratory are a couple of examples. Proteins, pathways, or both that are changed by the experimental treatments (Boyce et al., 2006; Daniels et al., 2006; Lippolis et al., 2006a) yield targets for future research. These approaches can be used to examine host-pathogen interactions to identify key bacterial or host proteins important to disease progression and recovery, organ development or host responses to stress. Proteins that are significantly altered in their expression, location, or are posttranslationally modified in animals with a disease compared with a group of healthy individuals represent protein targets for diagnostic tests, new therapies, or biomarker discovery.

Proteomics in Animal Health and Disease

The innate immune system is the first line of defense against a disease insult. The stresses of parturition

A

B
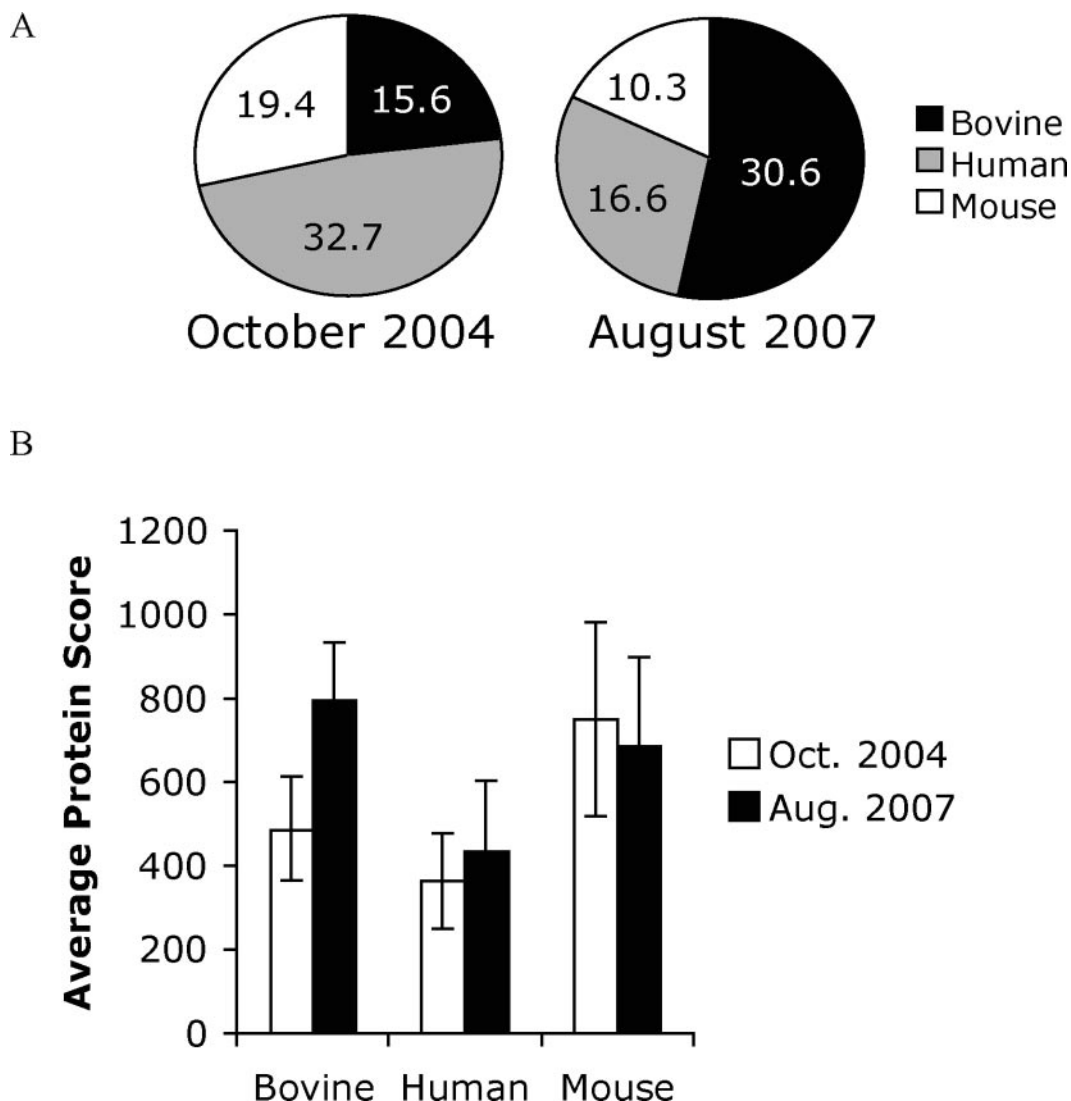
Figure 4. Effect of database of proteomic data quality. A data set of mass spectrometry data analyzing bovine neutrophil proteome was run against 2 protein databases. The first database was the National Center for Biotechnology Information's nonredundant protein database dated October 2004, and the second database was dated August 2007. Protein searches were run on the same day using the MASCOT (Matrix Science, Boston, MA) software with identical settings. (A) The percentage of the proteins identified as bovine (*Bos taurus*), human (*Homo sapiens*), or mouse (*Mus musculus*) was determined for each group. (B) The average MASCOT protein score was determined for all bovine, human, or mouse proteins.

and shipping have been clearly shown to suppress the innate immune system in cattle and calves (Kehrli et al., 1999; Duff and Galyean, 2007; Salak-Johnson and McGlone, 2007). The research demonstrating stress-induced immunosuppression has been accumulated in a large number of detailed experiments. A single proteomics experiment allows an investigator to examine stress models globally, in a search for new or unrecognized innate immunology pathways that are affected by stress in cattle.

The most common model for studying the effects of stress in cattle is glucocorticoid-induced stress (Roth and Kaeberle, 1982; Burton and Kehrli, 1995; Salak-Johnson and McGlone, 2007). Two recent proteomic experiments demonstrate the power of proteomics to quickly broaden our knowledge in this well-studied

area of animal immunology (Lippolis et al., 2006a; Mitchell et al., 2007). These proteomic papers identified new proteins/pathways that may be important in future hypothesis-driven studies of glucocorticoid-induced immunosuppression. More importantly, these papers raise important questions about the use of glucocorticoid-induced immunosuppression as a model of real-life animal stress. The work of Mitchell and coworkers demonstrated that glucocorticoids induced proteins in bronchoalveolar lavage fluid that could increase the susceptibility of calves to respiratory disease. However, they also found proteins that may enhance resistance to respiratory disease. The work of Lippolis and coworkers goes further in questioning the appropriateness of the glucocorticoid-induced immunosuppression model. Lippolis and coworkers

used quantitative shotgun proteomics to compare real periparturient-induced immunosuppression in dairy cow neutrophils to a glucocorticoid model of neutrophil immunosuppression. These studies are of particular importance because neutrophils are a key cell in the host response to mastitis infections. Their data clearly showed that there are both similarities and differences in bovine neutrophil protein expression in a naturally occurring immunosuppression observed in periparturient cows compared with a glucocorticoid-induced immunosuppression (Lippolis et al., 2006a). Therefore, this global proteomic approach to the study of bovine innate immunology has in one experiment raised a significant note of caution about a primary model of bovine immunosuppression as well as identifying new proteins that may be functionally significant in bovine innate immunity.

One of the new bovine innate immunity proteins found in this work was histone H2A.1 that was shown to be significantly downregulated in neutrophil membranes from immunosuppressed cows. The observation that H2A.1 was associated with membranes and was downregulated in immunosuppressed cows was a surprise of the type that can be found by proteomics. On the surface, the data were confusing, but a literature search turned up a reason for this observation and led to hypothesis-driven research on bovine neutrophil function. Recent work had shown a novel neutrophil-killing mechanism (Brinkmann et al., 2004). It involves the release by neutrophils of a web of DNA, histones, and antibacterial proteins, which forms a trap to kill bacteria. These complexes are referred to as neutrophil extracellular traps (**NET**). The histone data from the proteome analysis (Lippolis et al., 2006a) led to experiments that showed that this new neutrophil killing mechanism, NET, is present in bovine neutrophils (Lippolis et al., 2006b).

Proteomic-derived data on neutrophils will be of great importance in understanding many aspects of innate immune protection in diseases such as mastitis. Furthermore, the milk proteome has yielded evidence for the presence of previously unrecognized host defense proteins in milk (Smolenski et al., 2007) that may be important to mastitis research. Interestingly, proteomic examination of the milk fat globule membrane proteome (**MFGM**) for lactation biology research revealed the presence in the MFGM of the innate immune signaling molecules, TLR2, TLR4, and CD14 (Reinhardt and Lippolis, 2006). Because MFGM is derived from the apical membrane of the secretory cell, these new data suggest that mammary epithelium may be directly involved in signaling that may lead to early recruitment of innate immune cells to fight an infection. Additional quantitative proteomic studies have shown that TLR2 is significantly upregulated on MFGM in early lactation (Reinhardt and Lippolis, 2008), and at the same time these MFGM quantitative proteomic studies contributed new data on mammary development in early lactation.

The above discussion has focused on some recent proteomic contributions to host factors in animal health. Proteomic analysis of disease-producing bacteria also has great potential to increase our overall understanding of animal health. Bacterial proteomes compared between laboratory and disease states will increase our understanding of virulence factor expression, will lead to development of new diagnostic tests from new protein markers, and has the potential to identify new candidate proteins for vaccine development.

Several animal health laboratories are using proteomics for just these purposes. For example, Johne's disease (*Mycobacterium avium* subsp. paratuberculosis) is an economically significant disease in cattle. Proteomic experiments have been performed on this bacterium to identify new diagnostic tools (Cho et al., 2006) and to examine differences in the proteome between laboratory-adapted strains and clinical isolates (Radosevich et al., 2007) to better understand this disease for future Johne's disease diagnostics and vaccine development. Similar proteomic efforts and approaches are underway for *Brucella abortus* (Connolly et al., 2006), *Pasteurella multocida* (Boyce et al., 2006), *Mycobacterium bovis* (Mollenkopf et al., 2004), as well as other economically important production animal diseases.

Dairy Food Production Bacterial Proteomes. A large number of bacterial species are used in the production of dairy products. The proteins and pathways that are affected when these bacteria are moved from laboratory media to growth in milk are significant to the manufacturing process and are economically important. A proteomic approach can yield information regarding the global changes in protein expression needed to understand the adaptations that these bacteria undergo. The proteomes of *Lactobacillus bulgaricus, lactococcus lactis,* and *Streptococcus thermophilus* have been examined for protein changes resulting from growth in laboratory medium or in milk (Rechinger et al., 2000; Derzelle et al., 2005; Gitton et al., 2005). The proteomes of all 3 of these bacteria were significantly changed when the bacteria were grown in milk compared with laboratory media. These data provide food scientists with new information to modify/select these bacteria for optimum dairy food production practices. Moreover, these data suggest that we may get a more complete picture of bacterial pathogenesis in mastitis by examining changes in the bacterial proteome of common mastitis causing bacteria when grown in milk.

Muscle Biology Proteomes

Meat quality, appearance, flavor, and tenderness are affected by many factors (Bouley et al., 2004; Jia et al., 2006a, 2007). Postmortem storage times/temperature contribute significantly to meat characteristics in some as yet unknown ways. Proteomic approaches are being used extensively to examine postmortem changes in slaughtered beef and pork to understand the mecha-

nisms that control postmortem muscle changes that affect meat quality (Jia et al., 2006a,b, 2007; Sayd et al., 2006; Laville et al., 2007; Suman et al., 2007). As these findings are studied and extended experimentally, new and improved postmortem procedures may yield higher quality meat for consumers.

Proteomics in Animal Reproduction

High reproductive performance and factors that affect it have a significant impact on animal production profitability. Proteomics has been used to show that variation in protein types and amounts in seminal fluid regulates fertility indexes in dairy bulls (Kelly et al., 2006; Moura et al., 2006a,b). Whether these findings can be used to improve dairy bull fertility is as yet unknown, but these data would have been difficult to generate in the preproteomic era.

SUMMARY

We have briefly reviewed the significant technology, challenges, and promises of proteomic research. In addition, we have provided a snapshot of recent proteomic research in several areas of the animal sciences. It is clear that this is a science and technology that is in its infancy but a technology that shows great promise in leading to future advancements in animal science research. With the known potential for a disconnect between gene expression and protein expression, proteomic experiments in combination with gene expression experiments (e.g., microarray) will yield the most accurate and complete picture of the cellular response to environmental changes. A shotgun approach to protein expression has the advantage of being able to identify changes without prior knowledge of which proteins may change. Thus, proteomics has the ability to discover potentially new and unexpected connections between a change in treatment and a change in protein expression.

Challenges in the field of animal science proteomics include an incomplete genomic sequencing of some species and an incomplete understanding of protein function in all species. Genome sequencing programs are enabling molecular understanding of various species, whereas determination of the functions of many proteins will likely take many years. Proteomics is likely to play an important role in determining functions for many proteins. The promise of proteomics is that it will enable researchers to understand the complex interplay of potentially hundreds of proteins involved in a singular cellular process and to identify and quantitate these complex protein networks as a cell response to its environment.

LITERATURE CITED

Bouley, J., C. Chambon, and B. Picard. 2004. Mapping of bovine skeletal muscle proteins using two-dimensional gel electrophoresis and mass spectrometry. Proteomics 4:1811–1824.

Boyce, J. D., P. A. Cullen, V. Nguyen, I. Wilkie, and B. Adler. 2006. Analysis of the *Pasteurella multocida* outer membrane sub-proteome and its response to the in vivo environment of the natural host. Proteomics 6:870–880.

Brinkmann, V., U. Reichard, C. Goosmann, B. Fauler, Y. Uhlemann, D. S. Weiss, Y. Weinrauch, and A. Zychlinsky. 2004. Neutrophil extracellular traps kill bacteria. Science 303:1532–1535.

Burton, J. L., and M. E. Kehrli Jr. 1995. Regulation of neutrophil adhesion molecules and shedding of *Staphylococcus aureus* in milk of cortisol- and dexamethasone-treated cows. Am. J. Vet. Res. 56:997–1006.

Cho, D., N. Sung, and M. T. Collins. 2006. Identification of proteins of potential diagnostic value for bovine paratuberculosis. Proteomics 6:5785–5794.

Cho, W. C. 2007. Proteomics technologies and challenges. Genomics Proteomics Bioinformatics 5:77–85.

Connolly, J. P., D. Comerci, T. G. Alefantis, A. Walz, M. Quan, R. Chafin, P. Grewal, C. V. Mujer, R. A. Ugalde, and V. G. DelVecchio. 2006. Proteomic analysis of *Brucella abortus* cell envelope and identification of immunogenic candidate proteins for vaccine development. Proteomics 6:3767–3780.

Cravatt, B. F., G. M. Simon, and J. R. Yates 3rd. 2007. The biological impact of mass-spectrometry-based proteomics. Nature 450:991–1000.

Daniels, K. M., K. E. Webb Jr, M. L. McGilliard, M. J. Meyer, M. E. Van Amburgh, and R. M. Akers. 2006. Effects of body weight and nutrition on mammary protein expression profiles in Holstein heifers. J. Dairy Sci. 89:4276–4288.

Derzelle, S., A. Bolotin, M. Y. Mistou, and F. Rul. 2005. Proteome analysis of *Streptococcus thermophilus* grown in milk reveals pyruvate formate-lyase as the major upregulated protein. Appl. Environ. Microbiol. 71:8597–8605.

Duff, G. C., and M. L. Galyean. 2007. Board-invited review: Recent advances in management of highly stressed, newly received feedlot cattle. J. Anim. Sci. 85:823–840.

Elrick, M. M., J. L. Walgren, M. D. Mitchell, and D. C. Thompson. 2006. Proteomics: Recent applications and new technologies. Basic Clin. Pharmacol. Toxicol. 98:432–441.

Gingras, A. C., M. Gstaiger, B. Raught, and R. Aebersold. 2007. Analysis of protein complexes using mass spectrometry. Nat. Rev. Mol. Cell Biol. 8:645–654.

Gitton, C., M. Meyrand, J. Wang, C. Caron, A. Trubuil, A. Guillot, and M. Y. Mistou. 2005. Proteomic signature of *Lactococcus lactis* NCDO763 cultivated in milk. Appl. Environ. Microbiol. 71:7152–7163.

Griffin, T. J., S. P. Gygi, T. Ideker, B. Rist, J. Eng, L. Hood, and R. Aebersold. 2002. Complementary profiling of gene expression at the transcriptome and proteome levels in *Saccharomyces cerevisiae*. Mol. Cell. Proteomics 1:323–333.

Gygi, S. P., B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, and R. Aebersold. 1999a. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat. Biotechnol. 17:994–999.

Gygi, S. P., Y. Rochon, B. R. Franza, and R. Aebersold. 1999b. Correlation between protein and mRNA abundance in yeast. Mol. Cell. Biol. 19:1720–1730.

Hunt, D. F., R. A. Henderson, J. Shabanowitz, K. Sakaguchi, H. Michel, N. Sevilir, A. L. Cox, E. Appella, and V. H. Engelhard. 1992. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. Science 255:1261–1263.

Ideker, T., V. Thorsson, J. A. Ranish, R. Christmas, J. Buhler, J. K. Eng, R. Bumgarner, D. R. Goodlett, R. Aebersold, and L. Hood. 2001. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science 292:929–934.

Jia, X., M. Ekman, H. Grove, E. M. Faergestad, L. Aass, K. I. Hildrum, and K. Hollung. 2007. Proteome changes in bovine longissimus thoracis muscle during the early postmortem storage period. J. Proteome Res. 6:2720–2731.

Jia, X., K. I. Hildrum, F. Westad, E. Kummen, L. Aass, and K. Hollung. 2006a. Changes in enzymes associated with energy me-

tabolism during the early post mortem period in longissimus thoracis bovine muscle analyzed by proteomics. J. Proteome Res. 5:1763–1769.

Jia, X., K. Hollung, M. Therkildsen, K. I. Hildrum, and E. Bendixen. 2006b. Proteome analysis of early post-mortem changes in two bovine muscle types: M. longissimus dorsi and M. semitendinosis. Proteomics 6:936–944.

Kehrli, M. E., J. L. Burton, B. J. Nonnecke, and E. K. Lee. 1999. Effects of stress on leukocyte trafficking and immune responses: Implications for vaccination. Adv. Vet. Med. 41:61–81.

Kelly, V. C., S. Kuy, D. J. Palmer, Z. Xu, S. R. Davis, and G. J. Cooper. 2006. Characterization of bovine seminal plasma by proteomics. Proteomics 6:5826–5833.

Laville, E., T. Sayd, C. Terlouw, C. Chambon, M. Damon, C. Larzul, P. Leroy, J. Glenisson, and P. Cherel. 2007. Comparison of sarcoplasmic proteomes between two groups of pig muscles selected for shear force of cooked meat. J. Agric. Food Chem. 55:5834–5841.

Lippolis, J. D., B. D. Peterson-Burch, and T. A. Reinhardt. 2006a. Differential expression analysis of proteins from neutrophils in the periparturient period and neutrophils from dexamethasone-treated dairy cows. Vet. Immunol. Immunopathol. 111:149–164.

Lippolis, J. D., and T. A. Reinhardt. 2005. Proteomic survey of bovine neutrophils. Vet. Immunol. Immunopathol. 103:53–65.

Lippolis, J. D., T. A. Reinhardt, J. P. Goff, and R. L. Horst. 2006b. Neutrophil extracellular trap formation by bovine neutrophils is not inhibited by milk. Vet. Immunol. Immunopathol. 113:248–255.

Lippolis, J. D., F. M. White, J. A. Marto, C. J. Luckey, T. N. Bullock, J. Shabanowitz, D. F. Hunt, and V. H. Engelhard. 2002. Analysis of MHC Class II antigen processing by quantitation of peptides that constitute nested sets. J. Immunol. 169:5089–5097.

Mitchell, G. B., M. E. Clark, and J. L. Caswell. 2007. Alterations in the bovine bronchoalveolar lavage proteome induced by dexamethasone. Vet. Immunol. Immunopathol. 118:283–293.

Mollenkopf, H. J., L. Grode, J. Mattow, M. Stein, P. Mann, B. Knapp, J. Ulmer, and S. H. Kaufmann. 2004. Application of mycobacterial proteomics to vaccine design: Improved protection by *Mycobacterium bovis* BCG prime-Rv3407 DNA boost vaccination against tuberculosis. Infect. Immun. 72:6471–6479.

Moura, A. A., D. A. Chapman, H. Koc, and G. J. Killian. 2006a. Proteins of the cauda epididymal fluid associated with fertility of mature dairy bulls. J. Androl. 27:534–541.

Moura, A. A., H. Koc, D. A. Chapman, and G. J. Killian. 2006b. Identification of proteins in the accessory sex gland fluid associated with fertility indexes of dairy bulls: A proteomic approach. J. Androl. 27:201–211.

Moyer, S. C., B. A. Budnik, J. L. Pittman, C. E. Costello, and P. B. O'Connor. 2003. Attomole peptide analysis by high-pressure matrix-assisted laser desorption/ionization Fourier transform mass spectrometry. Anal. Chem. 75:6449–6454.

Nepom, G. T., J. D. Lippolis, F. M. White, S. Masewicz, J. A. Marto, A. Herman, C. J. Luckey, B. Falk, J. Shabanowitz, D. F. Hunt, V. H. Engelhard, and B. S. Nepom. 2001. Identification and modulation of a naturally processed T cell epitope from the diabetes-associated autoantigen human glutamic acid decarboxylase 65 (hGAD65). Proc. Natl. Acad. Sci. USA 98:1763–1768.

Ness, S. A. 2007. Microarray analysis: Basic strategies for successful experiments. Mol. Biotechnol. 36:205–219.

Ong, S. E., and M. Mann. 2005. Mass spectrometry-based proteomics turns quantitative. Nat. Chem. Biol. 1:252–262.

Ptacek, J., and M. Snyder. 2006. Charging it up: Global analysis of protein phosphorylation. Trends Genet. 22:545–554.

Radosevich, T. J., T. A. Reinhardt, J. D. Lippolis, J. P. Bannantine, and J. R. Stabel. 2007. Proteome and differential expression analysis of membrane and cytosolic proteins from *Mycobacterium avium* subsp. paratuberculosis strains K-10 and 187. J. Bacteriol. 189:1109–1117.

Rechinger, K. B., H. Siegumfeldt, I. Svendsen, and M. Jakobsen. 2000. "Early" protein synthesis of *Lactobacillus delbrueckii* ssp. bulgaricus in milk revealed by [35S] methionine labeling and two-dimensional gel electrophoresis. Electrophoresis 21:2660–2669.

Reinhardt, T. A., and J. D. Lippolis. 2006. Bovine milk fat globule membrane proteome. J. Dairy Res. 73:406–416.

Reinhardt, T. A., and J. D. Lippolis. 2008. Developmental changes in the milk fat globule membrane proteome during the transition from colostrum to milk. J. Dairy Sci. 91:2307–2318.

Reisinger, V., and L. A. Eichacker. 2006. Analysis of membrane protein complexes by blue native PAGE. Proteomics 6(Suppl 2):6–15.

Ross, P. L., Y. N. Huang, J. N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, and D. J. Pappin. 2004. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. Mol. Cell. Proteomics 3:1154–1169.

Roth, J. A., and M. L. Kaeberle. 1982. Effect of glucocorticoids on the bovine immune system. J. Am. Vet. Med. Assoc. 180:894–901.

Ryncarz, R. E., A. B. Dietz, and M. E. Kehrli Jr. 1995. Recognition of leukochimerism during genotyping for bovine leukocyte adhesion deficiency (BLAD) by polymerase-chain-reaction-amplified DNA extracted from blood. J. Vet. Diagn. Invest. 7:569–572.

Salak-Johnson, J. L., and J. J. McGlone. 2007. Making sense of apparently conflicting data: Stress and immunity in swine and cattle. J. Anim. Sci. 85:E81–E88.

Sayd, T., M. Morzel, C. Chambon, M. Franck, P. Figwer, C. Larzul, P. Le Roy, G. Monin, P. Cherel, and E. Laville. 2006. Proteome analysis of the sarcoplasmic fraction of pig semimembranosus muscle: Implications on meat color development. J. Agric. Food Chem. 54:2732–2737.

Schirmer, E. C., and L. Gerace. 2005. The nuclear membrane proteome: Extending the envelope. Trends Biochem. Sci. 30:551–558.

Smolenski, G., S. Haines, F. Y. Kwan, J. Bond, V. Farr, S. R. Davis, K. Stelwagen, and T. T. Wheeler. 2007. Characterisation of host defence proteins in milk Using a proteomic approach. J. Proteome Res. 6:207–215.

Stasyk, T., and L. A. Huber. 2004. Zooming in: Fractionation strategies in proteomics. Proteomics 4:3704–3716.

Stoughton, R. B. 2005. Applications of DNA microarrays in biology. Annu. Rev. Biochem. 74:53–82.

Suman, S. P., C. Faustman, S. L. Stamer, and D. C. Liebler. 2007. Proteomics of lipid oxidation-induced oxidation of porcine and bovine oxymyoglobins. Proteomics 7:628–640.

Wall, R. J., A. M. Powell, M. J. Paape, D. E. Kerr, D. D. Bannerman, V. G. Pursel, K. D. Wells, N. Talbot, and H. W. Hawk. 2005. Genetically enhanced cows resist intramammary *Staphylococcus aureus* infection. Nat. Biotechnol. 23:445–451.

Warnock, D. E., E. Fahy, and S. W. Taylor. 2004. Identification of protein associations in organelles, using mass spectrometry-based proteomics. Mass Spectrom. Rev. 23:259–280.

Washburn, M. P., D. Wolters, and J. R. Yates 3rd. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat. Biotechnol. 19:242–247.

Witze, E. S., W. M. Old, K. A. Resing, and N. G. Ahn. 2007. Mapping protein post-translational modifications with mass spectrometry. Nat. Methods 4:798–806.

Wu, C. C., K. E. Howell, M. C. Neville, J. R. Yates 3rd, and J. L. McManaman. 2000. Proteomics reveal a link between the endoplasmic reticulum and lipid secretory mechanisms in mammary epithelial cells. Electrophoresis 21:3470–3482.

Yates, J. R. 3rd. 1998. Mass spectrometry and the age of the proteome. J. Mass Spectrom. 33:1–19.

References

This article cites 58 articles, 19 of which you can access for free at:
http://jas.fass.org/cgi/content/full/86/9/2430#BIBL